

What is Claimed:

1. A method for creating a language model for a speech recognition system to indicate characters, the method comprising:

for each word phrase of a list of word phrases, associating a character string of the word phrase and the word phrase with a context cue indicative of identifying the character string;  
building a language model as a function of the associated word phrases and character strings.

2. The method of claim 1 wherein the language model comprises a statistical language model.

3. The method of claim 2 wherein the language model comprises an N-gram language model.

4. The method of claim 2 wherein the language model comprises a context-free-grammar.

5. The method of claim 1 wherein associating includes building a corpus of associated character strings and word phrases, and context cues, and wherein building the language model includes accessing the corpus.

6. The method of claim 1 wherein associating includes associating a first character of each word phrase with the word phrase.

7. The method of claim 6 wherein associating includes associating another character of at least some of the word

phrases, other than the first character, with the corresponding word phrases.

8. The method of claim 7 wherein associating includes associating each character of at least some of the word phrases with the corresponding word phrases.

9. The method of claim 7 wherein associating includes associating each character of each word phrase with the corresponding word phrase.

10. The method of claim 1 and further comprising adjusting a probability score for each of the associated character strings and word phrases in the language model.

11. The method of claim 1 wherein associating includes forming a phrase comprising the character string of the word phrase, the word phrase and the context cue for each word phrase of the list of word phrases.

12. The method of claim 11 wherein the context cue is similar to "as in" in English.

13. The method of claim 11 wherein the context cue comprises 的 in Chinese.

14. The method of claim 11 wherein the context cue comprises の in Japanese.

15. The method of claim 1 wherein each of the word phrases is a single word.

16. The method of claim 15 wherein each of the character strings is a single character.

17. The method of claim 1 wherein each of the character strings is a single character.

18. A computer readable medium having instructions, which when executed by a processor perform a method for recognizing characters when spoken, the method comprising:

receiving input speech having a character string, a word phrase having the character string and a context cue;

outputting the character string as text without the word phrase and the context cue.

19. The computer readable medium of claim 18 and further comprising instructions for:

accessing a language model indicative of a plurality of phrases, each phrase having a character string, a word phrase having the character string and a context cue.

20. The computer readable medium of claim 19 wherein the language model is indicative of phrases consisting essentially of associated character strings, word phrases having the character strings and context cues.

21. The computer readable medium of claim 19 wherein outputting the character string includes outputting the character string as a function of recognizing the character string using the language model.

22. The computer readable medium of claim 21 wherein the language model comprises a statistical language model.

23. The computer readable medium of claim 22 wherein the language model comprises an N-gram language model.

24. The computer readable medium of claim 21 wherein outputting the character string includes outputting the character string as only a function of an N-gram of the received input speech.

25. The computer readable medium of claim 21 wherein outputting the character string includes outputting the character string as a function of a comparison of a recognized character string with a recognized word phrase.

26. The computer readable medium of claim 25 wherein when the recognized character string is not present in the recognized word phrase, the character string that is outputted is a character string of the recognized word phrase.

27. The computer readable medium of claim 21 wherein the language model comprises a context-free-grammar.

28. The computer readable medium of claim 18 wherein each of the word phrases is a single word.

29. The computer readable medium of claim 28 wherein each of the character strings is a single character.

30. The computer readable medium of claim 18 wherein each of the character strings is a single character.

31. A computer readable medium having instructions, which when executed by a processor, for recognizing character strings when spoken, the instructions comprising:

- a language model indicative of phrases consisting essentially of associated character strings, word phrases having the character strings and context cues; and
- a recognition module for receiving data indicative of input speech, accessing the language model and outputting a character string spoken by the user wherein the input speech includes a word phrase having the character string and a context cue.

32. The computer readable medium of claim 31 wherein the recognition module outputs only the character string.

33. The computer readable medium of claim 31 wherein the language model comprises a statistical language model.

34. The computer readable medium of claim 31 wherein the language model comprises an N-gram language model.

35. The computer readable medium of claim 31 wherein the language model comprises a context-free-grammar.

36. The computer readable medium of claim 31 wherein the recognition module outputs the character string as a function of a comparison of a recognized character string with a recognized word phrase.

37. The computer readable medium of claim 36 wherein when the recognized character string is not present in the recognized word phrase, the character string that is outputted is a character string of the recognized word phrase.

38. The computer readable medium of claim 31 wherein each of the word phrases is a single word.

39. The computer readable medium of claim 38 wherein each of the character strings is a single character.

40. The computer readable medium of claim 31 wherein each of the character strings is a single character.